

# Spatial correlations in bed load transport: evidence, importance, and modelling.

J. Heyman<sup>1</sup>, H.B. Ma<sup>2</sup>, F. Mettra<sup>1</sup> and C. Ancey<sup>1</sup>

**Abstract.** This article examines the spatial dynamics of bed load particles in water. We focus particularly on the fluctuations of particle activity, which is defined as the number of moving particles per unit bed length. Based on a stochastic model recently proposed by *Ancey and Heyman* [2014], we derive the second moment of particle activity analytically; that is the spatial correlation functions of particle activity. From these expressions, we show that large moving particle clusters can develop spatially. Also, we provide evidence that fluctuations of particle activity are scale-dependent. Two characteristic lengths emerge from the model: a saturation length  $\ell_{sat}$  describing the length needed for a perturbation in particle activity to relax to the homogeneous solution, and a correlation length  $\ell_c$  describing the typical size of moving particle clusters. A dimensionless Péclet number can also be defined according to the transport model. Three different experimental data sets are used to test the theoretical results. We show that the stochastic model describes spatial patterns of particle activity well at all scales. In particular, we show that  $\ell_c$  and  $\ell_{sat}$  may be relatively large compared to typical scales encountered in bed load experiments (grain diameter, water depth, bed form wavelength, flume length...) suggesting that the spatial fluctuations of particle activity have a non-negligible impact on the average transport process.

## 1. Introduction

Originating in the late 1930s with the seminal work of Hans Albert Einstein [Einstein, 1937, 1950], the probabilistic approach to bed load transport has had a surge of interest among the scientific community in recent years [Papanicolaou *et al.*, 2002; Jerolmack and Mohrig, 2005; Ancey *et al.*, 2006, 2008; Valyrakis *et al.*, 2010; Ancey, 2010; Furbish and Schmeeckle, 2013]. This revival has been combined with a substantial improvement in laboratory measurement techniques. In particular, the use of high-speed videos of particle motion together with powerful digital processing, has allowed for ground-breaking precision in the description of sediment particle dynamics [Böhm *et al.*, 2004; Radice *et al.*, 2009; Lajeunesse *et al.*, 2010; Roseberry *et al.*, 2012; Martin *et al.*, 2012].

These data allow for an improved understanding of the transport process and its fluctuations. Indeed, bed load transport rates are known to show fluctuations often larger than the mean [Drake *et al.*, 1988; Kuhnle and Southard, 1988; Hoey, 1992; Ancey *et al.*, 2006; Singh *et al.*, 2009]. The problem arising in any system exhibiting internal fluctuations is the calculation of consistent average values, or relationships, that can be used to describe its macroscopic behavior.

This paper is concerned with drawing possible links between the microscopic stochastic motion of bed load particles and macroscopic variables, such as the average bed load flux or the particle activity (the number of moving particles per unit bed length). In other words, the question we try

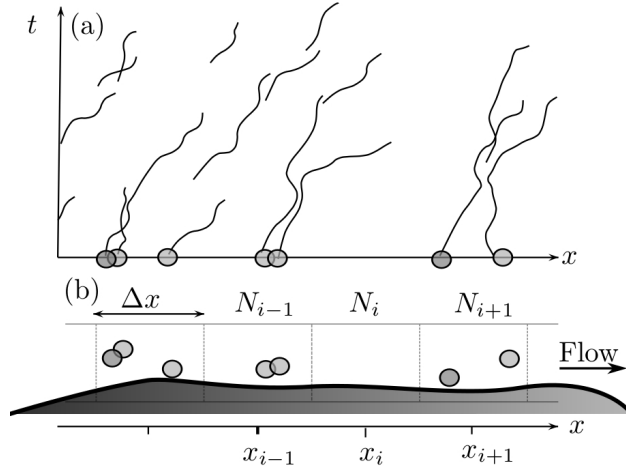
to answer here is how individual particle motion is reflected through larger scale transport relations. More precisely, how does noise, intrinsically present at small scales, modify the average macroscopic equilibrium at large scales? Until now, the variability of bed load flux has been deliberately ignored in most transport models. Is this approximation physically justified or, on the contrary, do models need to take into account bed load rate fluctuations in order to accurately predict and quantify sediment budgets?

Among the recent stochastic models of bed load transport, *Sun and Donahue* [2000] proposed a two states Markov model suggesting that bed load transport rates would follow a binomial distribution. *Wu and Chou* [2003] considered the rolling and lifting probabilities of particles in a turbulent stream while in *Wu and Yang* [2004], they proposed a stochastic partial transport model for mixed size sediments. *Turowski* [2010] suggested that the shape of the probability distribution of the bed load flux was a function of the inter-arrival time of particles. More recently, in four companion papers, *Furbish et al.* [2012] provided further insights into particle random motion and its consequences on macroscopic conservation equations. Using an ensemble averaging procedure, they found that the bed load flux comprises both an advective and a diffusive term due to particle velocity fluctuations. *Ancey et al.* [2008] developed a stochastic erosion/deposition model describing the fluctuation of the number of moving particles in an observation window. Using the framework of birth-death Markov processes, they provided a comprehensive picture of the large fluctuations observed in their experiments. Generalizing *Ancey et al.*'s [2008] probabilistic model, *Ancey and Heyman* [2014] were able to model the spatial variability of particle activity. By studying the erosion, deposition and motion of particles on a lattice made of regular cells, they ended up with a stochastic equation describing the process in both space and time. The model is valid for low to moderate transport rates.

In this paper, we explore some applications of the stochastic model recently proposed by *Ancey and Heyman* [2014]. While *Ancey and Heyman* [2014] paper concerned the theoretical foundations of the model, this article focuses on validation issues. In doing so, we will demonstrate how the

<sup>1</sup>Laboratory of Environmental Hydraulics, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne, Switzerland

<sup>2</sup>State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing, China



**Figure 1.** (a) Particle trajectories in a time-space plane. (b) Discretization of the space in cells of equal length ( $\Delta x$ ) and number of particles in each cell  $i$  at a given time  $t$  ( $N_i$ ).

model accurately reproduces spatial fluctuations of the bed load particle activity. This will be achieved by comparing theoretical results with various experimental data of particle trajectories in time and space (Fig. 1).

The paper is organized as follows. First, to make the article self-contained, we briefly go over how the general stochastic equations governing the bed load phase are derived. A rigorous and detailed derivation is not provided here since it is available in *Ancey and Heyman* [2014]. Then, we compute the second moment (the spatial correlation function) of the particle activity. We also give the analytical expression of the  $K$ -function [Ripley, 1976], often used in point process analysis to highlight the possible correlations existing between particle locations. In the last section, we use three different experimental studies to test the model, two of which have already been published [Böhm et al., 2004; Roseberry et al., 2012]. The third study is an original data set which consists of three experiments carried out in a steep slope flume. A general method to calibrate model parameters on experimental data is proposed. Finally, we discuss the importance of including spatial variability in bed load transport models and we propose possible improvements of *Ancey and Heyman*'s [2014] stochastic model.

## 2. Theory

### 2.1. Physical space

The transport of bed load particles occurs in a thin layer over the surface of an erodible bed. Particles generally move in a preferential direction (down the slope, parallel to fluid flow) so that it is possible to constrain the study to a one-dimensional space in that principal direction. A generalization to a two-dimensional space, while technically possible, goes beyond the scope of this paper.

Let us consider a one-dimensional space that represents a river reach, or an experimental flume. The space is divided into cells of equal length  $\Delta x$  (Fig. 1). Each cell of this lattice is labelled by an index  $i$ . We call  $N_i(t)$  the random variable describing the number of moving particles  $n_i$  in cell  $i$  at time  $t$ . We introduce the multivariate probability

$$P([n_1, n_2, \dots], t) = P(\mathbf{n}, t), \quad (1)$$

where  $\mathbf{n}$  is the vector of all  $n_i$ . In other words,  $P(\mathbf{n}, t)$  is the probability of simultaneously observing  $N_1(t) =$

$n_1, N_2(t) = n_2, \dots$  at time  $t$ . The particle activity in cell  $i$  is defined as

$$\gamma(x_i, t) = N_i(t)/\Delta x, \quad (2)$$

where  $x_i$  denotes the position of the center of the cell  $i$ .

### 2.2. Particle motion

Bedload transport describes the motion of bed particles (sliding, rolling or saltating) sheared by a fluid. To build their model, *Ancey et al.* [2008] and *Ancey and Heyman* [2014] distinguished three independent phases of particle motion: entrainment, transport and deposition. Those are briefly reviewed below.

The entrainment of a resting particle by a fluid flow has been extensively studied and its intermittent and random character is widely accepted [Einstein, 1950; Papanicolaou et al., 2002; Wu and Chou, 2003; Schmeeckle et al., 2007; Detert et al., 2010; Celik et al., 2010; Valyrakis et al., 2010; Dwivedi et al., 2011]. For flow conditions close to incipient sediment motion, the fluid flow intermittently dislodges particles from the bed. Turbulent flow structures being spatially and temporally correlated, it is also likely that several particles are entrained simultaneously, leading to clouds of moving particles [Nelson et al., 1995; Drake et al., 1988]. Various experiments suggested that different mechanisms of entrainment exist, such as entrainment caused by a particle collision or by a local bed rearrangement [Schmeeckle et al., 2001; Heyman et al., 2013].

In their model, *Ancey et al.* [2008] conceptualized the entrainment of a particle as a sum of two basic random processes: (i) a memoryless and uncorrelated process, referred to as entrainment, and (ii) a correlated process with intensity proportional to the number of particles already in motion, and referred to as collective entrainment. The probability of a particle being entrained ( $n_i \rightarrow n_i + 1$ ) in a cell  $i$  of length  $\Delta x$  during a small time interval  $dt$  is thus

$$P(n_i \rightarrow n_i + 1) = (\lambda \Delta x + \mu n_i) dt, \quad (3)$$

where  $\lambda$  [particles  $\text{m}^{-1} \text{s}^{-1}$ ] is the mean entrainment rate of particles per unit length and  $\mu$  [ $\text{s}^{-1}$ ] is the collective entrainment rate.

After being entrained, a particle is dragged by the fluid flow for a certain time before depositing onto the bed. *Ancey et al.* [2008] envisioned the deposition of a particle as a memoryless and independent random process. Thus, at any time, the probability of observing a single particle deposition ( $n_i \rightarrow n_i - 1$ ) in the cell  $i$  during  $dt$  can be expressed as

$$P(n_i \rightarrow n_i - 1) = \sigma n_i dt, \quad (4)$$

where  $\sigma$  [ $\text{s}^{-1}$ ] is the mean particle deposition rate. Note that, in this basis,  $1/\sigma$  is the mean travel time of a particle.

Once put in motion, particles are transported downstream by the fluid flow. Their velocity is frequently altered due to repeated impacts on the bed as well as by drag fluctuations due to turbulence. *Ancey and Heyman* [2014] proposed a model of Brownian motion in a potential to describe the transport process. They found that, under certain conditions, the transport of particles could be described locally by the sum of two contributions: (i) a deterministic advection at the average particle velocity  $\bar{u}_s$  and (ii) a random jump process between lattice cells that can be described by the following transition probabilities in the infinitesimal time interval  $dt$ :

$$\begin{aligned} P(n_i \rightarrow n_i - 1, n_{i-1} \rightarrow n_{i-1} + 1) &= d n_i dt, \\ P(n_i \rightarrow n_i - 1, n_{i+1} \rightarrow n_{i+1} + 1) &= d n_i dt, \end{aligned} \quad (5)$$

where  $d$  [ $\text{s}^{-1}$ ] is a local diffusivity rate. In the following statistical analysis, we do not consider the contribution of

advection in the transport of particles since it is deterministic and is not contributing to the fluctuations of the number of moving particles [Ancey and Heyman, 2014]. Naturally, the advective contribution to the particle transport process will be reintroduced later in the deterministic part of the equations.

### 2.3. Birth-death process and Poisson representation

In the previous section, we implicitly assumed that the rate coefficients  $\lambda, \sigma, \mu$  and  $d$  were constant in space and time. This leads us to focus on fluctuations that precisely originate from the randomness of particle motions and exchanges with the bed rather than fluctuations arising because of local changes in flow or bed slope—that would in turn modify the rate coefficients, when bedforms are present for instance.

The transition probabilities defined above form the elementary rules governing the evolution of a multivariate birth-death Markov process (i.e.: a memoryless process of many variables, which evolves by unitary jumps). From these simple rules, Ancey and Heyman [2014] derived the multivariate master equation describing the temporal evolution of  $P(\mathbf{n}, t)$ ,

$$\begin{aligned} \frac{\partial P(\mathbf{n}, t)}{\partial t} = & \sum_i d(n_i + 1) (P(\mathbf{n} + \mathbf{r}_i^+ + \mathbf{r}_{i+1}^-, t) + P(\mathbf{n} + \mathbf{r}_i^+ + \mathbf{r}_{i-1}^-, t)) \\ & + (n_i + 1)\sigma P(\mathbf{n} + \mathbf{r}_i^+, t) + (\lambda\Delta x + (n_i - 1)\mu) P(\mathbf{n} + \mathbf{r}_i^-, t) \\ & - (\lambda\Delta x + n_i(\sigma + \mu) + 2dn_i) P(\mathbf{n}, t), \end{aligned} \quad (10)$$

where  $\mathbf{r}_i^\pm$  is a vector whose elements are all zeros except for its  $i$ -th value:  $r_i = \pm 1$ ,  $r_k = 0$  for  $k \neq i$ .  $P(\mathbf{n} + \mathbf{r}_i^+ + \mathbf{r}_{i-1}^-, t)$  is thus the probability of observing the system in the state  $\mathbf{n}' = (n_1, n_2, \dots, n_{i-1} - 1, n_i + 1, n_{i+1}, \dots)$ .

This master equation can be greatly simplified using the Poisson representation. Similarly to Laplace or Fourier transforms in the spectral theory of time series, the Poisson representation is a linear operator that transforms a discrete probability space into a continuous one [Gardiner and Chaturvedi, 1977]. More precisely, it assumes that the probability function of  $N_i$  can be decomposed into Poisson distributions with various rates  $a_i$ ,

$$P(n_i, t) = \int_{\mathcal{R}^+} \frac{e^{-a_i} a_i^n}{n!} f(a_i, t) da_i. \quad (7)$$

Since the only parameter of a Poisson distribution is also its mean,  $a_i$  can be interpreted as the mean number of particles in the cell  $i$ . On the other hand,  $f(a_i, t)$  is the probability of observing the Poisson rate  $a_i$  in cell  $i$  at time  $t$ .

By inserting Eq. (1) in the master equation (6), Ancey and Heyman [2014] showed that  $f(\mathbf{a}, t)$ — $\mathbf{a}$  being the vector of all  $a_i$ —follows an explicit Fokker–Planck equation (i.e., a partial differential equation governing the time evolution of probability functions)

$$\begin{aligned} \frac{\partial}{\partial t} f(\mathbf{a}, t) = & \sum_i \mu \frac{\partial^2 a_i f(\mathbf{a}, t)}{\partial a_i^2} \\ & + \frac{\partial}{\partial a_i} [f(\mathbf{a}, t) (\lambda\Delta x - a_i(\sigma - \mu))] \\ & + \frac{\partial}{\partial a_i} [f(\mathbf{a}, t) d(a_{i+1} + a_{i-1} - 2a_i)]. \end{aligned} \quad (8)$$

Equivalently,  $a_i$  can be shown to follow a Langevin stochastic equation (i.e., a differential equation with both a deterministic and a stochastic parts)

$$\begin{aligned} da_i(t) = & (d(a_{i+1} + a_{i-1} - 2a_i) + \lambda\Delta x - a_i(\sigma - \mu)) dt \\ & + \sqrt{2\mu a_i} dW_i(t), \end{aligned} \quad (9)$$

where  $dW_i(t)$  is the derivative of a Wiener random process, which may be interpreted as a time uncorrelated noise (also called white noise). This noise is said to be multiplicative since its intensity is modulated by  $\sqrt{a_i}$  (in contrast to an additive noise which is independent of the state of the process, as it was for example assumed in Jerolmack and Mohrig [2005]).

Just like the definition of the particle activity  $\gamma(x, t)$ , let us call  $\eta(x, t)$  the Poisson rate per unit bed length (referred to as Poisson activity in the following). We have

$$\eta(x_i, t) = a_i(t)/\Delta x. \quad (10)$$

Using Eq. (9) and letting  $\Delta x \rightarrow 0$ , we obtain the Langevin stochastic partial differential equation for the Poisson activity

$$\begin{aligned} d\eta(x, t) = & [D\nabla^2 \eta(x, t) + (\mu - \sigma)\eta(x, t) + \lambda] dt \\ & + \sqrt{2\mu\eta(x, t)} dW(x, t), \end{aligned} \quad (11)$$

where  $W(x, t)$  is now a spatial Wiener process satisfying

$$dW(x, t) dW(x', t) = \delta(x - x') dt. \quad (12)$$

According to Eq. (12), the multiplicative noise term  $\sqrt{2\mu\eta} dW$  arising in Eq. (11) is perfectly uncorrelated in space and time. We also introduced the notation  $D = d\Delta x^2$ , which highlights the connection between the local particle jump rate  $d$  [ $\text{s}^{-1}$ ] and the macroscopic particle diffusivity  $D$  [ $\text{m}^2 \text{s}^{-1}$ ].

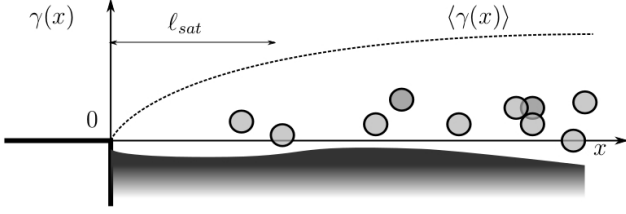
Making use of the linearity of the deterministic part of Eq. (11), we can reintroduce the deterministic advection flux

$$\begin{aligned} d\eta(x, t) = & [-\bar{u}_s \nabla \eta(x, t) + D\nabla^2 \eta(x, t)] dt \\ & + [\lambda - (\sigma - \mu)\eta(x, t)] dt \\ & + \sqrt{2\mu\eta(x, t)} dW(x, t) \end{aligned} \quad (13)$$

Eq. (13) models the stochastic evolution of the rate (per unit length) of the Poisson distribution followed by  $N_i(t)$ . It is shown in Appendix B how Eq. (13) can be solved numerically and how it can be related to the point process framework [Cox and Isham, 1980]. Eq. (13) also shares interesting similarities with the BCRE model of dry granular avalanches of Bouchaud et al. [1995]. In Appendix A, we show how Eq. (13) may be used as a stochastic version of the BCRE model in order to characterize spatial correlations in some dry granular flows.

### 2.4. Moments

In the remaining of the paper, the notation  $\langle \bullet \rangle$  denotes ensemble averaging (i.e., average over all the possible states of a stochastic process). There exists a simple connection between moments of  $a$  in the Poisson representation and moments of the real variable  $N$ . Indeed, the  $p$ -factorial moment of  $N$  (i.e.,  $\langle n(n-1)(n-2)\dots \rangle$ ) is equal to the  $p$ -moment of  $a$  (i.e.,  $\langle a^p \rangle$ ), implying that  $\langle n \rangle = \langle a \rangle$  and  $\langle n^2 \rangle = \langle a^2 \rangle + \langle a \rangle$  [Gardiner and Chaturvedi, 1977; Ancey and Heyman, 2014]. Similar relationships exist between between moments of  $\eta$  and moments of  $\gamma$ :  $\langle \eta(x, t) \rangle = \langle \gamma(x, t) \rangle$  and  $\langle \gamma(x, t) \gamma(x', t) \rangle = \langle \eta(x, t) \eta(x', t) \rangle + \delta(x - x') \langle \eta(x, t) \rangle$



**Figure 2.** Boundary value problem corresponding to the relaxation of particle activity to equilibrium.  $\ell_{sat}$  is the saturation length.

[Gardiner and Chaturvedi, 1977]. In the following, we study the first and second moments of Eq. (13).

#### 2.4.1. First moment

The average behavior of  $\eta(x, t)$  is easily obtained by dropping the noise term in Eq. (13),

$$\frac{\partial \langle \eta \rangle}{\partial t} + \bar{u}_s \frac{\partial \langle \eta \rangle}{\partial x} = D \frac{\partial^2 \langle \eta \rangle}{\partial x^2} + \lambda - (\sigma - \mu) \langle \eta \rangle. \quad (14)$$

It is a linear advection-diffusion-reaction equation. For  $t \rightarrow \infty$ , providing that  $\sigma > \mu$ , the stationary and homogeneous solution is

$$\langle \eta \rangle_s = \frac{\lambda}{\sigma - \mu}, \quad (15)$$

where the notation  $\langle \bullet \rangle_s$  denotes the ensemble average for stationary and homogeneous conditions. Thus, the stationary homogeneous particle activity is  $\langle \gamma \rangle_s = \lambda/(\sigma - \mu)$ .

Ancey and Heyman [2014] studied the evolution of a sediment pulse in space and time. Here, we focus on the stationary behavior of (14) given a fixed Dirichlet boundary condition at the origin  $\langle \gamma(0) \rangle = 0$  (Fig. 2). The problem simplifies to a second order ordinary differential equation whose solution is given by

$$\begin{aligned} \langle \gamma(x) \rangle &= \frac{\lambda}{\sigma - \mu} \left( 1 - e^{-x/\ell_{sat}} \right), \\ \ell_{sat} &= \frac{2\ell_c}{\text{Pe}} \left( \sqrt{1 + 4\text{Pe}^{-2}} - 1 \right)^{-1}, \end{aligned} \quad (16)$$

where we have introduced  $\ell_c = \sqrt{D/(\sigma - \mu)}$  and the dimensionless number  $\text{Pe} = \bar{u}_s \ell_c / D$ , which can be interpreted as a local Péclet number. The Péclet number is usually defined as the ratio of the rate of advection by the flow to the rate of diffusion. In the case of the diffusion of matter, it can also be defined as the product of a typical length scale by the advection velocity divided by the diffusivity. In our case, the local Péclet number compares particle diffusion against advection with respect to the correlation length  $\ell_c$ . The latter originates from the coupled action of diffusion and particle exchanges with the bed (collective entrainment and deposition) but its meaning will be best understood while studying spatial correlations. Note that, in contrast to aeolian sediment transport, the saturation length  $\ell_{sat}$  does not originate from particle inertia (which is negligibly small in water) but rather from the particle exchanges with the bed and their transport by the flow [Charru, 2006].

Unfortunately, no experimental data about saturation length in bed load transport under water are presently available, so that no comparison could be made with the theoretical predictions.

#### 2.4.2. Second moment

We show in Appendix C1 that the stationary and homogeneous spatial correlation function of the particle activity reads

$$\begin{aligned} \langle \gamma(x), \gamma(x') \rangle_s &\equiv \langle \gamma(x) \gamma(x') \rangle_s - \langle \gamma(x) \rangle_s^2 \\ &= \delta(x - x') \langle \gamma \rangle_s + \frac{\langle \gamma \rangle_s \mu}{2\ell_c(\sigma - \mu)} \exp\left(-\frac{|x - x'|}{\ell_c}\right), \end{aligned} \quad (17)$$

where we have already used the correlation length  $\ell_c = \sqrt{D/(\sigma - \mu)}$ . The steady-state spatial correlation function is thus the sum of a Dirac delta function (i.e.,  $\delta(x)$  is a distribution which is zero everywhere except at  $x = 0$ , with an integral of one over the entire real line) of intensity equal to the mean density of moving particles and an exponentially decaying function corresponding to correlations caused by the collective entrainment of particles (when  $\mu = 0$ , this term disappears). The correlation length  $\ell_c$  modulates the speed of the decay. Thus,  $\ell_c$  is a measure of the typical spatial scale of fluctuations in particle activity. It increases with the diffusivity of particles and with the collective entrainment rate, but decreases with the deposition rate. Interestingly enough, when  $\mu = 0$ , no more spatial correlations are observed in the particle activity although  $\ell_c$  remains positive. When  $\mu \rightarrow \sigma$ , the spatial correlations of particle activity become infinitely large and when  $\mu = \sigma$ , Eq. (13) becomes unstable and an exponential increase in the number of moving particles is observed.

Another quantity of interest, often used to describe a spatial point process, is the conditional intensity  $h(x - x')$ , which gives the conditional probability of finding a particle at  $x'$  given that there is a particle at  $x$  [Cox and Isham, 1980]. The conditional intensity and the correlation function are directly related by

$$\langle \gamma(x, t), \gamma(x', t) \rangle_s = \delta(x - x') \langle \gamma \rangle_s + \langle \gamma \rangle_s h(x - x') - \langle \gamma \rangle_s^2,$$

so that by identification, we have

$$h(x - x') = \langle \gamma \rangle_s + \frac{1}{2\ell_c} \frac{\mu}{\sigma - \mu} \exp\left(-\frac{|x - x'|}{\ell_c}\right). \quad (18)$$

A more convenient function for data analysis is the  $K$ -function [Ripley, 1976], where  $K(x)$  represents the expected number of moving particles found in a ball of radius  $x$  centered on a particle location, divided by the mean process rate. This can be calculated from the conditional intensity function by

$$\begin{aligned} K(x) &= \frac{1}{\langle \gamma \rangle_s} \int_0^x h(u) du \\ &= x + \frac{1}{\langle \gamma \rangle_s} \frac{\mu}{\sigma - \mu} \left[ 1 - \exp\left(-\frac{x}{\ell_c}\right) \right]. \end{aligned} \quad (19)$$

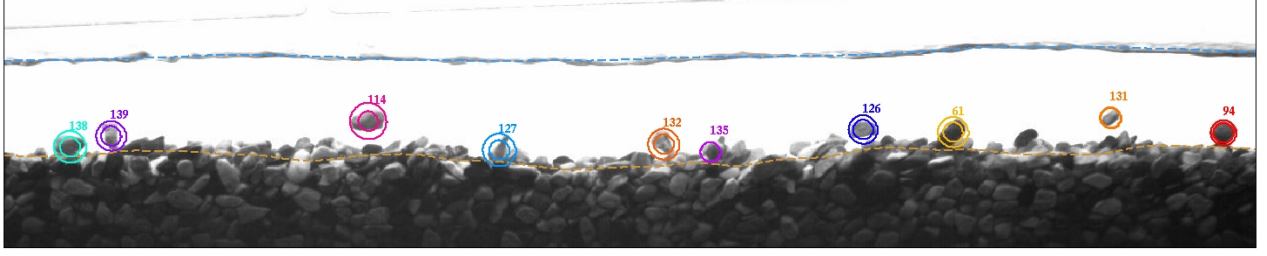
In the case of a Poisson point process in one dimension (i.e., a process which is spatially uncorrelated),  $K(x) = x$ . Furthermore, Eq. (19) shows that  $K(x) > x$  if  $\mu > 0$ ; so the point process formed by particle locations is said to be clustered (see Appendix B).

## 3. Applications

After deriving the spatial correlation function of the particle activity, we examine how it can be compared with experimental data.

### 3.1. Experiments

We use three different experimental data sets. Two of them have been previously published [Böhm et al., 2004; Roseberry et al., 2012]. The third comes from an experimental setup especially built by the authors to observe spatial



**Figure 3.** Snapshot from one of the two cameras used in experiments J (the field of vision is about 50 cm long). Moving particle locations are estimated by the tracking algorithm (circles show the confidence interval of the particle diameter estimates, while numbers refer to the trajectory index  $j$ ). The water surface and the bed elevation are also detected automatically (blue and orange dashed lines).

**Table 1.** Experimental parameters and model fits.  $B$  [cm], channel width;  $d_{50}$  [mm], mean particle diameter;  $\tau_s$  [-], Shields stress;  $Fr$  [-], Froude number;  $\tan(\theta)$  [%], slope angle;  $\bar{v}$  [m s<sup>-1</sup>], mean flow velocity;  $\bar{h}$  [cm], mean water depth;  $\bar{q}_s$  [particles s<sup>-1</sup>], mean output solid discharge;  $\bar{\gamma}$  [particles m<sup>-1</sup>], mean activity;  $\bar{u}_s$  [m s<sup>-1</sup>], average particle velocity;  $\sigma$  [s<sup>-1</sup>], deposition rate;  $D$  [cm<sup>2</sup> s<sup>-1</sup>], particle diffusivity;  $\lambda$  [particle m<sup>-1</sup> s<sup>-1</sup>], particle entrainment rate per meter length;  $\mu$  [s<sup>-1</sup>], collective entrainment;  $\ell_c$  [cm], correlation length ( $\ell_c = \sqrt{D/(\sigma - \mu)}$ );  $Pe$  [-] local Péclet number ( $Pe = \ell_c \bar{u}_s / D$ );  $I(\infty)$  [-], limiting value of the dispersion index;  $\ell_{sat}$  [cm] saturation length (Eq. (16)).

	$B$	$d_{50}$	$\tau_s$	$Fr$	$\tan\theta$	$\bar{v}$	$\bar{h}$	$\bar{q}_s$	$\bar{\gamma}$	$\bar{u}_s$	$\sigma$	$D$	$\lambda$	$\mu$	$\ell_c$	$Pe$	$I(\infty)$	$\ell_{sat}$
B10-5	0.6	6	0.11	1.42	10.0	0.41	1.0	4.9	26.9	0.170	2.72	15	24	1.825	4.1	4.6	3.0	20
R0-79	6.0	0.5	0.06	0.35	-	0.31	12.5	78.7	1711.0	0.046	1.85	1.5	171	1.754	3.8	12.2	18.6	47
J3-1	3.5	7	0.14	1.30	3.5	0.80	3.8	1.4	4.6	0.310	0.52	59	0.33	0.447	28.7	15.1	7.2	434
J4-1	3.5	7	0.17	1.39	4.5	0.86	3.9	1.2	3.9	0.300	0.50	89	0.39	0.403	29.8	10.1	5.0	303
J5-1	3.5	7	0.14	1.47	4.7	0.80	3.1	0.9	2.8	0.320	0.56	55	0.27	0.470	24.3	14.1	6.0	345

and temporal fluctuations of bed load transport (Fig. 3). All three studies provide high resolution measurements of particle trajectories using high speed videos. Details of the experimental setups are given in Appendix D while experimental conditions are reported in Table 1.

Hereafter, we denote all *Böhm et al.* [2004] experiments by using the prefix B, *Roseberry et al.* [2012] experiment using R, and the new data set using J. The numbers following the prefix specify experimental slope and solid discharge. For instance B10-5 stands for *Böhm et al.* [2004] experiment conducted using a 10% sloping flume with a mean solid discharge of 5 particles s<sup>-1</sup>.

Typically, an experimental outcome consists of an ensemble of particle locations in the streamwise direction:  $x_{j,k}$  is the position of particle  $j$  at frame  $k$ . In total, experiment B gathers more than 8000 particle trajectories over 4 minutes, experiment R gathers more than 300 particles trajectories over about 0.4 s while each experiment J gathers in average 5000 trajectories over 10 minutes.

### 3.2. Spatial fluctuations

Let us consider the number of moving particles in a window of length  $L$  at a given time  $t$ ,

$$N(L, t) = \int_L \gamma(x, t) dx. \quad (20)$$

When  $t \rightarrow \infty$ , the stationary average of  $N(L, t)$  is

$$\begin{aligned} \text{Mean}[N(L)] &\equiv \left\langle \int_L \gamma(x, t) dx \right\rangle_s = \int_L \langle \gamma(x, t) \rangle_s dx \\ &= \langle \gamma \rangle_s L, \end{aligned} \quad (21)$$

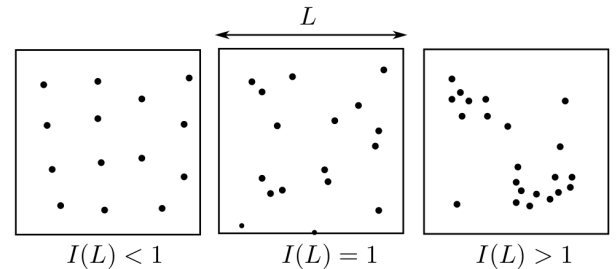
while the variance of  $N(L, t)$  (sometimes called the variance of the sample mean) is defined by

$$\begin{aligned} \text{Var}[N(L)] &\equiv \left\langle \int_L \gamma(x, t) dx \int_L \gamma(x', t) dx' \right\rangle_s - \left\langle \int_L \gamma(x, t) dx \right\rangle_s^2 \\ &= \int_L \int_L \langle \gamma(x, t) \gamma(x', t) \rangle_s dx dx'. \end{aligned} \quad (22)$$

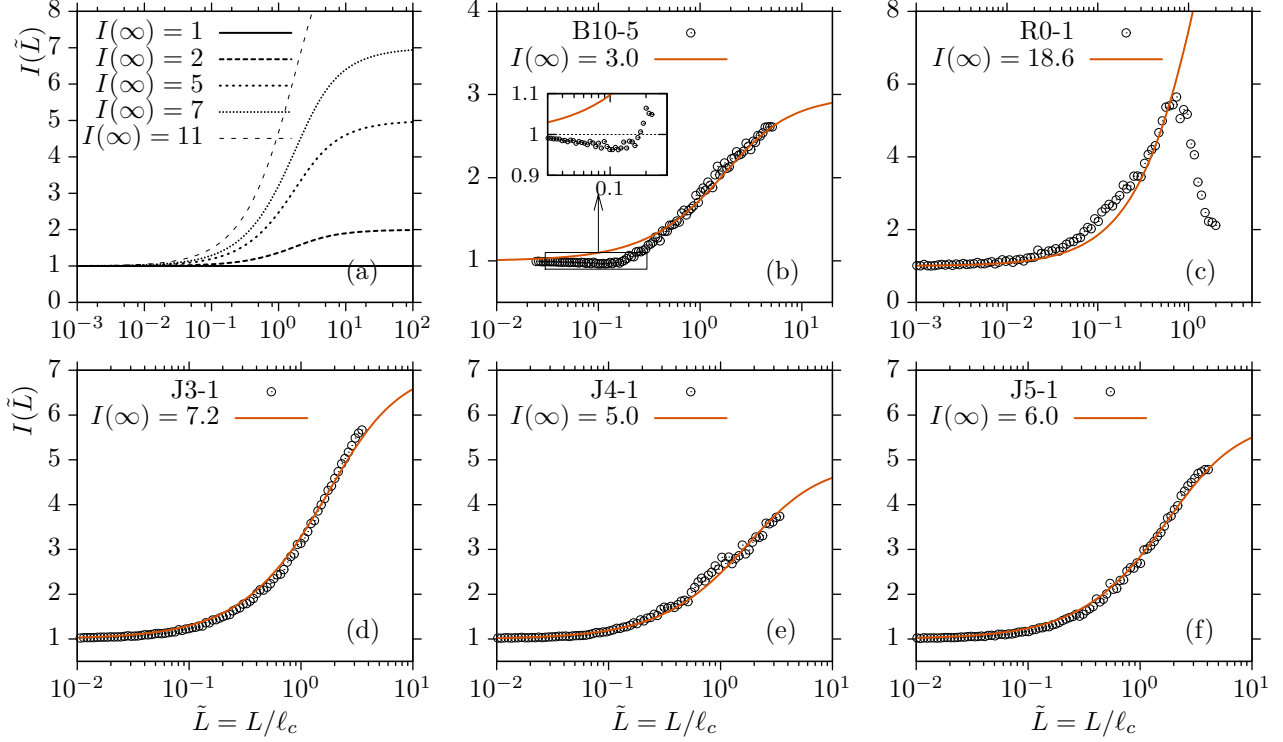
Introducing Eq. (17) into Eq. (22) and integrating it (see Appendix C2), we find

$$\begin{aligned} \text{Var}[N(L)] &= \langle \gamma \rangle_s L \\ &+ \langle \gamma \rangle_s \ell_c \frac{\mu}{\sigma - \mu} \left( L/\ell_c + e^{-L/\ell_c} - 1 \right). \end{aligned} \quad (23)$$

Eq. (23) shows the dependence of the variance of  $N(L)$  on the length  $L$  of the observation window. Let us define the dispersion index  $I(L)$  as the ratio of the variance over



**Figure 4.** Example of a realization of point positions in a two-dimensional space depending on the value of the dispersion index  $I(L)$ . Here,  $L$  can be interpreted as the size of the box.



**Figure 5.** Dispersion index ( $I(\tilde{L})$ ), with  $\tilde{L} = L/\ell_c$ . (a) Eq. (24), (b) experiment B, (c) experiment R, (d-e-f) experiments J.

the mean

$$I(\tilde{L}) = \frac{\text{Var}[N(\tilde{L})]}{\text{Mean}[N(\tilde{L})]} = 1 + \frac{\mu}{\sigma - \mu} \left( 1 + \frac{e^{-\tilde{L}} - 1}{\tilde{L}} \right), \quad (24)$$

with  $\tilde{L} = L/\ell_c$ .

The dispersion index is used to characterize the relative positions of points (particle locations). Three classes of stochastic processes are generally distinguished depending on the value of  $I$ : under-dispersed processes for  $I < 1$ ; purely random processes (or Poisson processes) when  $I = 1$ ; and over-dispersed or clustered processes when  $I > 1$  (Fig. 4).

The theoretical dispersion index (24) grows from one, when the observation window is small, to the constant value  $I(\infty) = 1 + \mu/(\sigma - \mu)$ , as the window length tends to infinity (Fig. 5(a)). In other words, depending on the observation scale  $L$ , the number of moving particles in the observation window exhibits a different statistical behavior.

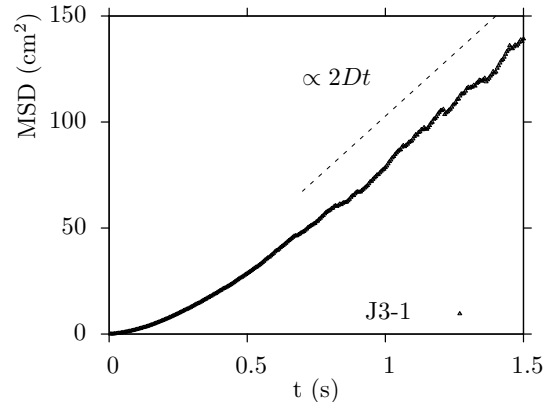
When  $L$  tends to 0,  $I(L)$  tends to one, so that the variance and the mean of  $N(L)$  are equal. Thus, in the small scale limit,  $N(L)$  tends to a Poisson process. This is mathematically explained by the presence of the Dirac delta function in the spatial correlation function (17). In other words, for decreasing values of  $L$ , a limit will be reached when most of the time an observation window contains no particle, and more rarely one. The small  $L$  limit can be seen as representative of a Bernoulli process (i.e.,  $N(L) = 1$  with probability  $\langle \gamma \rangle_s L$  and  $N(L) = 0$  with probability  $1 - \langle \gamma \rangle_s L$ ), which is well approximated by a Poisson process when  $\langle \gamma \rangle_s L \rightarrow 0$ .

On the contrary, when  $L \rightarrow \infty$ ,  $I$  reaches a constant value  $I(\infty)$ . Note that  $I(\infty) > 1$  if  $\mu > 0$ , so that the variance of  $N(L)$  is now greater than its mean. Thus, for larger scales and when  $\mu > 0$ ,  $N(L)$  cannot be described anymore by a Poisson process. Thus, moving particles are expected to form clusters during their motion when collective entrainment is considered (Fig. 4).

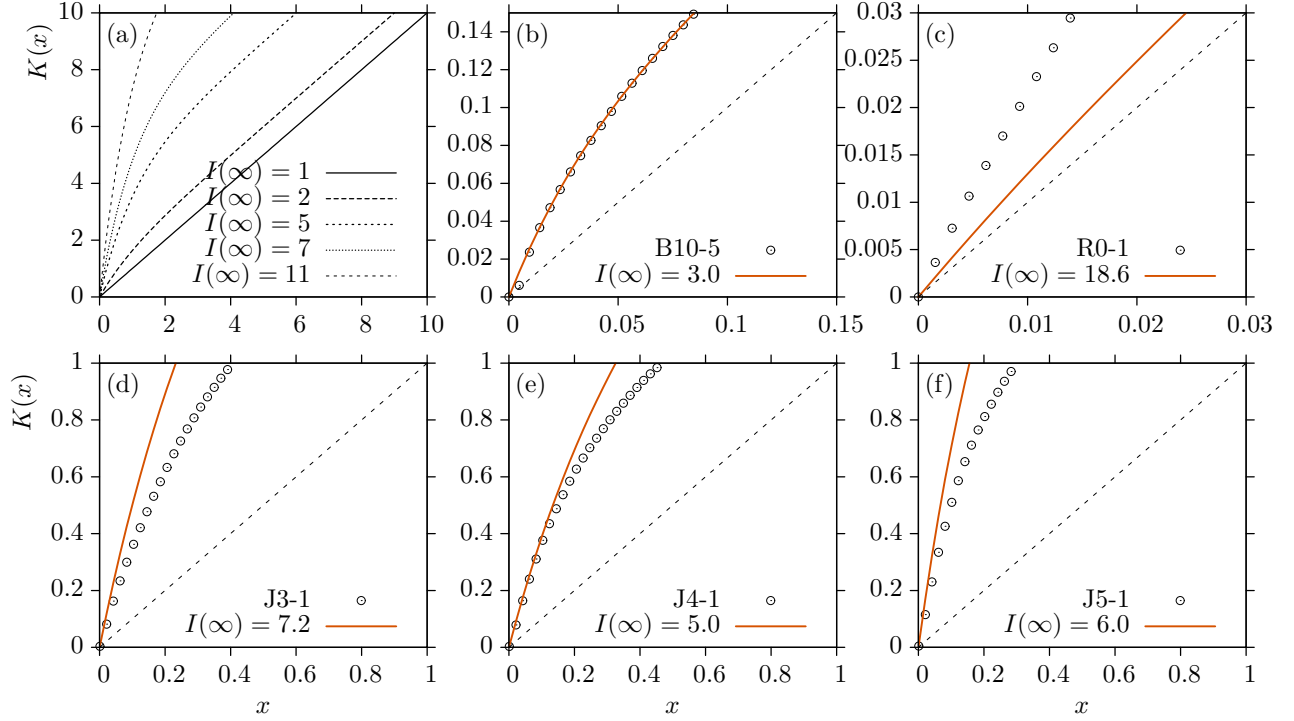
Experimental dispersion indices are presented in Fig. 5 (b-f). The procedure used to compute such index is pre-

sented in Appendix D4. In all experiments, dispersion indices change through spatial scales. From a Poisson type process at small scales ( $\tilde{L} \rightarrow 0$ ),  $I(\tilde{L})$  increases with increasing scales (Fig. 5). The dispersion index of experiment R follows a slightly different evolution, since it drops at scales larger than 5 cm. This behavior might be explained by the relatively short measurement window ( $\sim 8$  cm) and the relatively small number of frames available in experiment R, which leads to biased estimates of the dispersion index (see Appendix D4).

One striking feature of the experimental dispersion index that appears only in experiment B is the slight decrease below unity for lengths of the order of the particle diameter (Fig. 5(b)). This phenomenon results from negative values in the correlation function at those scales and cannot be de-



**Figure 6.** Mean squared displacement (MSD) of particles for experiment J3-1. The dashed line stands for the linear diffusion limit.



**Figure 7.**  $K$ -function. (a) Eq. (19) for  $\langle \gamma \rangle_s = 1$  and  $\ell_c = 1$ , (b) experiment B, (c) experiment R, (d-e-f) experiments J. The dashed line corresponds to the Poissonian case ( $K(x) = x$ ). Note: fits were calculated from the dispersion index results.

scribed by the Markov model. Indeed the theoretical spatial correlation function (17) is strictly greater than zero so that the dispersion index is expected to grow monotonically. The presence of negative values in the experimental correlation function is explained by the finite diameter of particles. Experiments B took place in a one-dimensional channel whose width equals particle diameters. Thus, there is less probability of finding two particles separated by a distance smaller than the particle diameter, resulting in negative correlation at the diameter scale.

Unfortunately, in none of the experiments presented, the limiting value of the dispersion index  $I(\infty)$  is reached for the maximum measurement length. Even experiments J — designed specially to achieve this purpose — are unable to reach the final plateau. It is still possible to extrapolate the theoretical expression (24) to predict the behavior of fluctuations at larger scales and to obtain  $I(\infty)$ .

Experimental  $K$ -functions are displayed in Fig. 7(b-f). In Appendix D4, we precise how the  $K$ -function can be estimated experimentally. In all experiments,  $K(x)$  is greater than  $x$ , suggesting that correlations exist between particle locations.

### 3.3. Parameter estimates

In the following, we show how the parameters  $\bar{u}_s$ ,  $D$ ,  $\lambda$ ,  $\sigma$  and  $\mu$  can be estimated given the experimental particle trajectories.

While the average particle velocity  $\bar{u}_s$  is simply determined by the arithmetic mean of all instantaneous particle velocities, the diffusivity  $D$  can be obtained by computing the experimental mean squared displacement of moving particles. Indeed, the mean squared displacement of particles obeying pure diffusion increases linearly with time:  $\langle x(t) - \langle x(t) \rangle \rangle^2 = 2Dt$  [Taylor, 1922]. As particle motions are correlated, the mean squared displacement is not linear at small time, but tends to the linear diffusion limit at large time [Uhlenbeck and Ornstein, 1930]. Thus, the limiting diffusivity of particles is obtained by fitting a line to the asymptotic mean squared displacement of particles (Fig. 6).

Note that, the mean squared displacement is only computed with the parts of trajectories where the particle moves, excluding the rest periods.

The mean deposition rate of particles  $\sigma$  is readily obtained by counting deposition events in a given observation window during a given time, then dividing this number by the observation time and by the mean number of moving particles in the window.

$\mu$  and  $\lambda$  are the only parameters which cannot be estimated independently. The theoretical expressions of first and second moments have to be used to determine their respective values. First, we adjust  $\mu$  so that Eq. (24) matches the experimental dispersion index. Then,  $\lambda$  is simply obtained using equation (15). It is worth highlighting that only two parameters ( $\lambda$  and  $\mu$ ) are tuned to fit first and second moments, while the three others are estimated independently. Eq. (24) captures extremely well the experimental dispersion indices at all scales (Fig. 5(b-f)). Estimated parameters are reported in Table 1.

Based on these estimates, we compare the theoretical  $K$ -function (19) to each experiment. From Fig. 7(b-f), we can see that the theoretical  $K$ -function describes experimental data less accurately than the dispersion index does. This could be explained by the fact that an additional parameter, the mean particle activity  $\langle \gamma \rangle_s$ , is required in Eq. (19). Thus, if the measure of  $\langle \gamma \rangle_s$  is biased (as it can be the case for short trajectory samples or for non-homogeneous transport conditions), the agreement between theory and experiments is also biased. Another explanation may be that a two-dimensional  $K$ -function would be more appropriate for two-dimensional experiments (such as experiment R).

### 3.4. Characteristic lengths and local Péclet number

Once  $D$ ,  $\sigma$  and  $\mu$  have been determined,  $\ell_c$  can be calculated. Looking at equations (17) and (24),  $\ell_c$  defines the scale at which approximately 36.8% of the maximum fluctuations above the mean are observed. In other words, to



observe at least 95% of the total fluctuations, lengths of the order of  $20\ell_c$  should be observed experimentally, that is, about 1 m for experiments B and R, and 3 m for experiments J. In addition to the technological challenge such a long acquisition length involves, the experimental flume might need to be even longer to avoid effects from the input and output boundary conditions.

The saturation length  $\ell_{sat}$  (see Eq. (16)) is a good estimate of the length needed for perturbations induced at the boundary to dissipate. It ranges from tens of centimeters in experiments B and R to several meters for experiments J (Table 1). Unfortunately, no experimental data about saturation length in bed load transport under water are presently available, so that no comparison could be made with these predictions. In any case, it is clear that the boundary conditions of experiments B and J, carried out in relatively short flumes (between 0.5 and  $10\ell_{sat}$ ), may have an influence on the results.

For experimental lengths of the order of  $20\ell_c$ , it is often impossible to insure the spatial homogeneity of sediment transport. Indeed, the instability of the bed-water interface leads to the development of bedforms of various wavelengths (from centimeters to hundreds of meters) and thus precludes the use of constant parameters.

The local Péclet number can be calculated with the obtained values of parameters (Table 1).  $Pe$  is observed to range from 4 (experiment B) to 14 (experiments R and J), showing the variety of modes of transport of bed load. Bed load occurring in experiment B is still strongly diffusive at the correlation length scale while in experiments R and J, it is mostly advective at this scale.

## 4. Summary and discussion

In this paper, we studied the spatial fluctuations of the number of moving particles per unit bed length, also called the particle activity [Furbish *et al.*, 2012]. These fluctuations have been shown to have a great deal of effect on the measurements of bed load transport rates in both field and experimental surveys [Gomez *et al.*, 1990; Dinehart, 1992; Garcia *et al.*, 2000; Cudden and Hoey, 2003; Bunte and Abt, 2005]. The model recently proposed by Ancey and Heyman [2014], generalizing the probabilistic model of Ancey *et al.* [2008] to a spatial dimension, offers a simple theoretical framework to understand and quantify these fluctuations.

The stochastic model is based on five parameters, most of which have a physical meaning: the entrainment rate  $\lambda$ , the collective entrainment rate  $\mu$ , the deposition rate  $\sigma$ , the average particle velocity  $\bar{u}_s$  and the particle diffusivity  $D$ . The first two have to be estimated via the method of moments (with Eqs. (15) and (22)) while the three lasts can be calibrated independently. Two characteristic lengths emerge from the model: a saturation length  $\ell_{sat}$  (Eq. (16)) quantifying the length needed for particle activity to recover its average equilibrium value, and a correlation length  $\ell_c = \sqrt{D/(\sigma - \mu)}$  which describe the typical size of fluctuations in particle activity. We also defined a local Péclet number  $Pe = \bar{u}_s \ell_c / D$  that describes the relative importance of advection against diffusion of particles at the correlation length. This number plays an important role in the value of the saturation length.

The stochastic model was tested against various experimental data of particle trajectories, and showed good overall agreement, notably in the description of the dispersion index.

This study also provides interesting guidelines for researchers studying the fluctuations of bed load transport rates. To capture 95% of the fluctuations of particle activity, an experiment should be designed such that it provides

a measurement window larger than  $20\ell_c$ . The difficulty lies in the fact that  $\ell_c$  is not known *a priori*, but it has to be computed after parameters estimation or measured directly if the limiting value of the dispersion index could be reached experimentally.

Another issue arises since, at lengths of the order of  $20\ell_c$ , it is generally difficult to ensure that bed load transport is homogeneous. Indeed, bedforms inexorably develop and migrate, modifying locally the flow and sediment transport. Model parameters may thus vary in time and space, precluding the use of the preceding results, derived in the case of stationary and homogeneous transport conditions.

More generally, this leads us to question the use of average equations (such as Eq. (14)) to describe bed load transport. As fluctuations were shown to span over scales often larger than the ones at which bed load can be considered stationary and homogeneous—or even at scales larger than the experiment size—average equations may fail at describing the non-linear interactions that may exist between the fluctuations in particle activity and the changes in bed elevation and water velocity for instance. Consequently, a correct description of bed load transport cannot avoid the modelling of local fluctuations and their interactions with the system boundaries.

Imagine now that the stochastic sediment transport equation (13) is coupled with Exner and Saint-Venant or Navier-Stokes equations. By a simultaneous and local description of the bed load activity fluctuations, as well as the fluid flow and the bed surface evolution, we may be able to give a more accurate picture of the whole transport process.

The proposed model may also be generalized to a second spatial dimension. Indeed, the motion of particles is only rarely unidirectional, as particle collisions and turbulent flow drag tend to modify particle trajectories [Seizilles *et al.*, 2014]. Thus, a dispersion of particles in the direction normal to the mean sediment velocity vector may also occur. The corresponding cross-stream diffusivity is expected to be different than the streamwise diffusivity, so that the overall diffusion process might be anisotropic. Still, owing to the overall linearity of the equations, the addition of a cross-stream diffusion term is straightforward.

## Notation

- $\langle \bullet \rangle$  Ensemble average.
- $\langle \bullet \rangle_s$  Ensemble average in steady state and homogeneous conditions.
- $\langle \bullet, \bullet \rangle$  Covariance of two random variables. For instance,  $\langle X, X \rangle = \langle X^2 \rangle - \langle X \rangle^2$ .
- $N_i/n_i$  Random variable/number of moving particles in cell  $i$ .
- $\Delta x$  Cell length in m.
- $\gamma(x, t)$  Density of moving particles at location  $x$  and time  $t$  in particles  $m^{-1}$ .
- $\mathbf{n}$  Vector of the number of moving particles in each cell.
- $x_i$  Position of the center of the cell  $i$ .
- $\lambda$  Average particle entrainment rate per meter length in particles  $m^{-1} s^{-1}$ .
- $\mu$  Average collective entrainment rate in  $s^{-1}$ .
- $\sigma$  Average particle deposition rate in  $s^{-1}$ .
- $D$  Macroscopic diffusivity in  $m^2 s^{-1}$ .
- $\bar{u}_s$  Mean particle velocity in  $m s^{-1}$ .
- $d$  Local diffusivity in  $s^{-1}$ .
- $\mathbf{r}_i^\pm$  Vector whose all but one value are zero:  $r_i = \pm 1$ ,  $r_k = 0$  for  $k \neq i$ .
- $a_i$  Poisson rate in cell  $i$  in the Poisson representation of  $n_i$ .



$\mathbf{a}$	Vector of Poisson rates in each cell.
$f$	Pseudo-density function of $\mathbf{a}$ .
$dW_i(t)$	derivative of a temporal Wiener process (white noise).
$\eta(x, t)$	Poisson density of moving particles at location $x$ and time $t$ in particles $\text{m}^{-1}$ .
$dW(x, t)$	derivative of a spatio-temporal Wiener process (two-dimensional white noise).
$\langle \gamma \rangle_s$	Steady-state homogeneous average density of moving particles in particles $\text{m}^{-1}$ .
$\langle \eta \rangle_s$	Steady-state homogeneous average Poisson density of moving particles in particles $\text{m}^{-1}$ .
$\ell_c$	Correlation length in m.
$\ell_{sat}$	Saturation length in m.
Pe	Local Péclet number (dimensionless number).
$I$	Index of dispersion.
$L$	Length of the observation window in m.
$\tilde{L}$	dimensionless length ( $L/\ell_c$ ) of the observation window.
$K$	$K$ -function.

**Acknowledgments.** This work was supported by the Swiss National Science Foundation under grant number 200021\_129538 (a project called “The Stochastic Torrent: stochastic model for bed load transport on steep slope”), the R’Equip grant number 206021\_133831 and by the competence center in Mobile Information and Communication Systems (grant number 5005-67322, MICS project).

## Appendix A: Link to the BCRE model

The BCRE model presented by *Bouchaud et al.* [1995] gives the density of rolling grains  $\mathcal{R}$  as the solution of:

$$\partial_t \mathcal{R} + \nabla(V\mathcal{R}) = \nabla^2(D\mathcal{R}) - \mathcal{R}\alpha\nabla h, \quad (\text{A1})$$

where  $\nabla h$  stands for the bed slope variations close to the angle of repose and  $\alpha$  is a constant. Thus, in their model, when the slope is bigger than the angle of repose ( $\nabla h < 0$ ) the second term on the left-hand side acts as a source in the equation. In that case, the number of rolling grains increases exponentially, leading to a local avalanche. On the contrary, when the slope is less than the angle of repose ( $\nabla h > 0$ ), grains are mainly deposited, causing the avalanche to stop ( $\mathcal{R} = 0$ ). The resemblance with Eq. (13) is striking. In the latter, an exponential increase in the number moving particles occurs when the collective entrainment rate is greater than or equal to the deposition rate ( $\mu \geq \sigma$ ). In contrast to (A1), when deposition is greater than collective entrainment, a non trivial steady-state solution exists, due to the uncorrelated particle entrainment process (with rate  $\lambda$ ).

Our model could thus be seen as a “BCRE” model that includes an additional random perturbation. Though the present work concerns bed load transport, and we restrict ourselves to the steady-state case ( $\mu < \sigma$ ), the limit  $\mu \rightarrow \sigma$  might be of particular interest for other granular flows. In particular, we suggest that (13) may also be applicable to certain dry and dilute granular flow, and thus may allow for their statistical description.

## Appendix B: Link to point processes

It is possible to draw an analogy between equation (13), obtained in the framework of birth-death Markov processes through Poisson representation, and the point process framework. Indeed, point processes are often defined by their rate function  $\eta(x, t)$  [Cox and Isham, 1980]. The simplest case is when the rate function is constant in time and space, resulting in a Poisson point process. When the rate function is a function of space and/or time, the process is called an inhomogeneous Poisson point process. Eventually, when the rate function is also a random variable, the process is called a doubly stochastic process, or Cox process [Cox and Isham, 1980]. This is the case with Eq. (13). To summarize, starting from a multivariate Markov process defined on lattice cells and described by a master equation, we end up with a model belonging to a general class of point processes, called doubly stochastic processes.

We now show how it is possible to simulate a probable realization of particle positions from Eq. (13). As noted earlier, by means of the Poisson representation,  $\eta(x, t)$  can be interpreted as the random rate of a Poisson distribution. First, we need to compute Eq. (13), to get a realization of  $\eta(x)$  at a given time  $t$ . This can be achieved using standard methods for stochastic differential equations (for instance an Euler–Maruyama scheme, which is an explicit finite-difference numerical scheme for stochastic equations [Kloeden and Platen, 2011]). Once we get a realization of  $\eta(x)$ , we proceed as follows. We choose a constant  $C > \eta(x)$  and compute a realization of point positions according to a Poisson process with rate  $C$ . This can be achieved by drawing  $CL$  random point locations from the uniform distribution between 0 and  $L$ ,  $L$  being the length of the computation area. Then, each point is randomly selected or discarded according to the criteria:

if  $r < \eta(x_k)/C$ , keep point;  
if  $r > \eta(x_k)/C$ , delete point;

where  $r$  is drawn from a uniform distribution in  $[0, 1]$ . The remaining points form a possible observation of particle positions according to the model (Fig. 8).

In Fig. 8, it is possible to observe the clustering of particles around the region of high  $\eta(x)$  values, while for the Poisson process, particles positions are purely random so no clustering appears. The clustering of particles is a special feature of our model (when  $\mu > 0$ ) and can be quantified by the study of the second moment.

## Appendix C: Theoretical developments

### C1. Spatial correlation function

Let  $g(x, x', t)$  denote the spatial correlation function of the Poisson density variable  $\eta(x, t)$ . By definition we have

$$\begin{aligned} g(x, x', t) &= \langle \eta(x, t) \eta(x', t) \rangle \\ &= \langle \eta(x, t) \eta(x', t) \rangle - \langle \eta(x, t) \rangle \langle \eta(x', t) \rangle. \end{aligned} \quad (\text{C1})$$

Taking the differential of  $g$  and using Itô’s calculus rules (an equivalent of the chain rule for stochastic equations),

$$\begin{aligned} dg(x, x', t) &= d \langle \eta(x, t) \eta(x', t) \rangle \\ &= \langle d\eta(x, t) \eta(x', t) \rangle + \langle \eta(x, t) d\eta(x', t) \rangle \\ &\quad + \langle d\eta(x, t) d\eta(x', t) \rangle. \end{aligned} \quad (\text{C2})$$

Note that  $d \langle \eta(x, t) \rangle \langle \eta(x', t) \rangle$  is zero by definition of the average. It comes

$$\begin{aligned} dg(x, x', t) &= D \left( \partial^2 / \partial x^2 + \partial^2 / \partial x'^2 \right) \langle \eta(x, t) \eta(x', t) \rangle dt \\ &\quad - \bar{u}_s \left( \partial / \partial x + \partial / \partial x' \right) \langle \eta(x, t) \eta(x', t) \rangle dt \\ &\quad - 2(\sigma - \mu) \langle \eta(x, t) \eta(x', t) \rangle dt + 2 \langle \eta \rangle_s (\lambda + \mu \delta(x - x')) dt. \end{aligned} \quad (\text{C3})$$

In a spatially homogeneous situation,  $g(x, x', t)$  is a function of  $r = |x - x'|$  only, which we call  $g(r, t)$ . Thus, substituting Eq. (C1) into Eq. (C3), we obtain

$$\frac{1}{2} \frac{\partial g(r, t)}{\partial t} = D \frac{\partial^2 g(r, t)}{\partial r^2} - (\sigma - \mu) g(r, t) + \mu \langle \gamma \rangle_s \delta(r). \quad (\text{C4})$$

The advection term disappears because  $\partial / \partial x = -\partial / \partial x'$ . Thus, the spatial correlation has no dependence on the mean velocity of particles. We look for the stationary behavior of (C4), which is

$$D \frac{\partial^2 g_s(r)}{\partial r^2} - (\sigma - \mu) g_s(r) + \mu \langle \gamma \rangle_s \delta(r) = 0, \quad (\text{C5})$$

with  $r = |x - x'|$ . We can simplify Eq. (C5) by rescaling the variable  $r$  by  $\tilde{r} = r / \ell_c$  where  $\ell_c = \sqrt{D / (\sigma - \mu)}$ . It yields

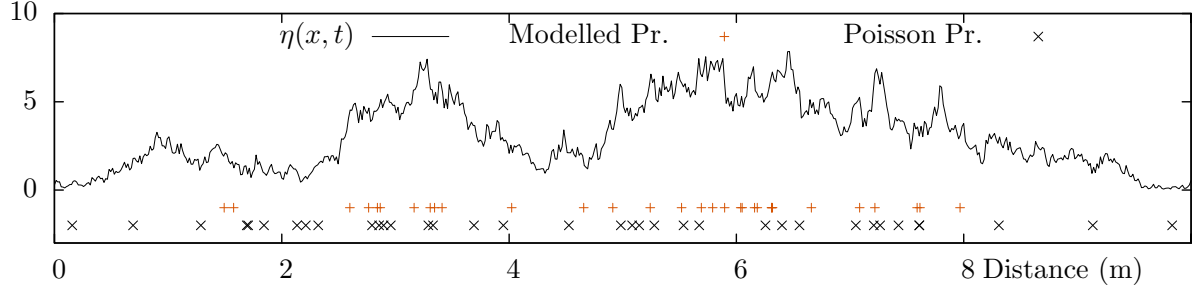
$$\frac{\partial^2 g_s(\tilde{r})}{\partial \tilde{r}^2} - g_s(\tilde{r}) + \frac{\langle \gamma \rangle_s}{\ell_c} \frac{\mu}{\sigma - \mu} \delta(\tilde{r}) = 0. \quad (\text{C6})$$

By means of Fourier transform, we obtain the algebraic equation

$$G(\omega) = \frac{\langle \gamma \rangle_s}{\ell_c} \frac{\mu}{\sigma - \mu} \frac{1}{\omega^2 + 1}, \quad (\text{C7})$$

where  $G(\omega)$  is the Fourier transform of  $g_s(\tilde{r})$ . The Fourier inverse of Eq. (C7), is given by

$$g_s(\tilde{r}) = \frac{\langle \gamma \rangle_s}{2\ell_c} \frac{\mu}{\sigma - \mu} \exp(-|\tilde{r}|). \quad (\text{C8})$$



**Figure 8.** Example simulation of the rate process (Eq. (13)) and corresponding possible realization of particle positions. We also plot the Poissonian case (with the same mean rate) for comparison. Model parameters are  $\lambda = 0.05$  particles  $\text{m}^{-1}$ ,  $\mu = 9.99 \text{ s}^{-1}$ ,  $\sigma = 10 \text{ s}^{-1}$ ,  $\bar{u}_s = 0.1 \text{ m s}^{-1}$  and  $D = 0.008 \text{ m}^2 \text{ s}^{-1}$

Hence

$$\langle \eta(x), \eta(x') \rangle_s = \frac{\langle \gamma \rangle_s \mu}{2\ell_c \sigma - \mu} \exp\left(-\frac{|x - x'|}{\ell_c}\right). \quad (\text{C9})$$

Second moment of  $\eta$  and  $\gamma$  are connected through the simple relationship [Gardiner and Chaturvedi, 1977]

$$\langle \eta(x, t), \eta(x', t) \rangle = \langle \gamma(x, t), \gamma(x', t) \rangle - \delta(x - x') \langle \gamma(x, t) \rangle,$$

so that Eq. (17) is recovered.

## C2. Spatial fluctuations

We wish to compute the integral

$$\text{Var}[N(L)] = \int_L \int_L \langle \gamma(x, t), \gamma(x', t) \rangle_s dx dx'.$$

That is

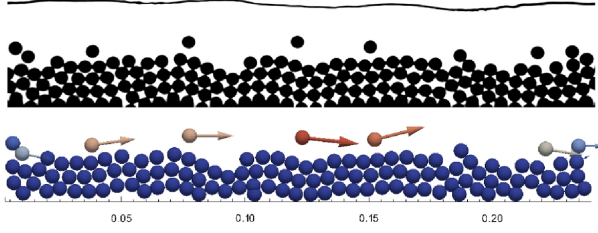
$$\begin{aligned} \text{Var}[N(L)] &= \langle \gamma \rangle_s L \\ &+ \frac{\langle \gamma \rangle_s \mu}{2\ell_c \sigma - \mu} \int_{-L/2}^{L/2} \int_{-L/2}^{L/2} e^{-|x-x'|/\ell_c} dx dx'. \end{aligned}$$

The value of the integral can be obtained by using

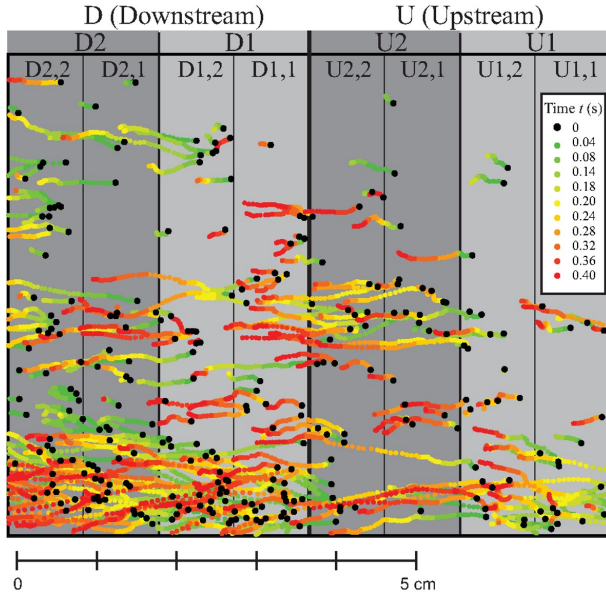
$$\begin{aligned} &\int_{-L/2}^{L/2} \int_{-L/2}^{L/2} e^{|x-x'|/\ell_c} dx dx' = \\ &\int_{-L/2}^{L/2} \left[ \int_{-L/2}^x e^{(x-x')/\ell_c} dx + \int_x^{L/2} e^{-(x-x')/\ell_c} dx \right] dx' = \\ &\ell_c \int_{-L/2}^{L/2} \left[ 2 - e^{L/(2\ell_c)} \left( e^{-x'/\ell_c} + e^{x'/\ell_c} \right) \right] dx' = \\ &2\ell_c^2 \left( L/\ell_c + e^{-L/\ell_c} - 1 \right). \end{aligned}$$

Thus

$$\text{Var}[N(L)] = \langle \gamma \rangle_s L + \langle \gamma \rangle_s \ell_c \frac{\mu}{\sigma - \mu} \left( L/\ell_c + e^{-L/\ell_c} - 1 \right).$$



**Figure 9.** Image recorded during a B experiment (top) and visualization of particles velocity vectors after image processing (bottom). Units are in meter.



**Figure 10.** Map view of experiments R showing particle motions occurring during the 0.4 sec time series; note the clustering of motions, partly reflecting effects of the turbulent sweeps. (Reproduced from *Roseberry et al.* [2012] with the authorization of the authors and AGU.)

## Appendix D: Experiments

### D1. Experiments B

These experiments were carried out in a narrow steep flume where sediment consisted of glass beads of equal size (6 mm). Particle transport was completely two-dimensional; this allowed *Böhm et al.* [2004] to take pictures through the side wall and detect and track individual particles via image processing. Camera resolution was  $640 \times 192$  pixels with a frame rate of 129.2 frame per seconds (fps). Each sequence comprised 8000 images corresponding to a duration of approximately 1 min. The acquisition length was 22.5 cm, for a resolution of 0.3 mm/pixel. Thus this imaging technique covers about 2 orders of magnitude in space. For further information on the experimental conditions, the reader is referred to [*Böhm et al.*, 2004; *Ancey et al.*, 2008].

Fig. 9 shows an example of a recorded image and the corresponding reconstruction of particle positions and velocities using image processing.

### D2. Experiment R

*Roseberry et al.* [2012] presented a set of experiments where particle trajectories were sampled in a two-

dimensional window of the bed viewed from the top. High-speed imaging at 250 fps over a 7.57 cm (streamwise) by 6.05 cm (cross-stream) bed-surface domain, and with  $1280 \times 1024$  pixels resolution provided the basis for tracking particle motions (with a precision of 0.06 mm/pixel). Bed material consisted of relatively uniform coarse sand with an average diameter of  $d_{50} = 0.5$  mm.

The data set involved one experiment with a total duration of 0.4 seconds, i.e. 100 frames (Fig. 10). In contrast to the three other data sets, experiment R concerns relatively small particles (sand) over mild slope (the slope is not given in [*Roseberry et al.*, 2012] but the Froude number is much lower than unity).

### D3. Experiments J

The originality of this data set compared to the two others lies in its high temporal and spatial resolutions. Two cameras of  $1280 \times 200$  pixels resolution, placed side by side, took pictures from the transparent side wall at a rate of 200 fps. The length of the observation window was slightly less than 1 m (with a precision of about 0.4 mm/pixel) while the duration of a sequence was 150 seconds (30 000 images). For each experiment, four film sequences were repetitively taken to insure good statistical results.

Experiments were carried out in a 2.5-m-long flume. The erodible bed was made of natural sediment particles with mean diameter of 8 mm. The flume was 3.5 cm wide and the water depth was ranging from 3 to 4 cm during experiments. The channel slope ranged between 3% and 5%. The flow was fully supercritical. Small antidunes were occasionally growing and propagated upstream, but the bed remained nearly flat in all experiments. As the channel width to flow depth ratio was relatively small ( $B/h \sim 1$ ), the fraction of the shear stress taken by the bed was certainly reduced, because of the increased side wall friction. Experimental studies [*Knight*, 1981] report a drop of about 40 to 60% of the bed shear stress for such aspect ratio. It is thus hard to determine precisely the experimental Shield stress without any direct flow velocity measurements.

Image processing and automatic particle tracking were then performed on these images (Fig. 3). The processing steps from raw images to particle trajectories were the following:

1. First the raw images were processed using the powerful yet simple method of median background subtraction [*Yilmaz et al.*, 2006; *Radice et al.*, 2006]. This allows a distinction between an immobile background (made up of particle resting on the bed) and a moving foreground (the moving particles).

2. An algorithm was then used to detect the centroid position of the moving particles in the foreground images. This was achieved after thresholding the foreground image and computing properties of connected regions (such as area, barycentre, eccentricity...)

3. Moving particles between two consecutive images were then associated into trajectories. The Hungarian algorithm was used here to obtain the best combinations. In case of conflict (for instance, if two particles are assigned to the same particle in the following frame), the trajectory was supposed to end and a new trajectory was built.

4. Finally, to reconstruct broken trajectories, a Kalman filter was applied to each missing measurements and overlapping trajectories were merged.

### D4. Experimental dispersion index and $K$ -function

The experimental dispersion index at length  $L$  is estimated as follows. First, we randomly choose  $p$  sub-windows of equal length  $L$  inside the whole available observation

region. As the sub-windows are selected randomly, they possibly overlap. Then, we construct the vector  $N_{i,k}$  of the number of moving particles found in the sub-window  $i$  ( $i = 1, \dots, p$ ) at frame  $k$  ( $k = 1, \dots, f$ ,  $f$  being the total number of video frames).  $\text{Var}[N(L)]$  and  $\text{Mean}[N(L)]$  are then estimated with all  $N_{i,k}$  samples. Typically, we chose  $p = 20$ . In experiments J, the dispersion index is thus estimated over more than 2 millions of samples. Note that, such a large number of samples is necessary to get unbiased estimates of moments, since samples may not be independent of each other. Similarly, the dispersion index of experiment R is estimated over 2000 samples and the dispersion index of experiment B is estimated over 640 000 samples.

The procedure to compute the experimental  $K$ -function is given in *Ripley* [1976]. In the latter paper, several methods are presented to prevent the problems arising at the boundaries of the image. Indeed, for a particle located close to the image boundary, the number of particles found in a circle of radius  $x$  larger than the particle-boundary distance may be underestimated. To prevent this, we chose to limit the computation of  $K(x)$  to particles located at a minimum distance  $x$  from the image boundary.

